Supplementary Material

S1

Domains	**Amount**
KS	394
AT	539
DH	261
ER	87
KR	457
ACP	609
TE	29

Table S1. Number of experimentally characterized domains retrieved from the DoBISCUIT database (release version 2014-03-10).

S2

$$max_s\left(\Sigma_{i=1}^{as}\frac{AA_{i,s}}{AT_s}\right)$$

Formula S2. Position probability matrices to predict substrate specificity of AT domains

s = substrates: malonyl, methylmalonyl, ethylmalonyl, methoxymalonyl;
as = active site residues: Q643, L671, Y742, S744 and L795 of DEBS AT domain;
AT = AT domains;
AA = 20 natural amino acids: A, C, E, D, G, F, I, H, K, M, L, N, Q, P, S, R, T, W, V, Y.

Scores	Evaluation	Points
Length score	Length of the molecule, based on the main chain	1 point for each matching segment
		-1 point for each not matching segment
Main bond score	Bond type of each segment	1 point for each matching bond
		-1 point for each not matching bond
Main atom score	Atom type of each segment	1 point for each matching atom
		-1 point for each not matching atom
Neighbor score	Compares the neighbors for each segment, the type of neighbor is not evaluated in this score.	+1 if a neighbor exists in both segments
		+1 if a neighbor does not exists in either segment
		-1 if one neighbor exists in only one segment
		-2 if 2 neighbors exist in only one segment
		0 if one segment has 2 neighbors and the other only 1
For simplification purposes the following three scores are combined as ‘side chain score’ in the server output.		
Side score	Compares the neighbors for each segment, very strict both atom and bond must match. Can only match ones.	+1 if matched
2nd neighbor score	Compares the next neighbor, the type of neighbor is not evaluated in this score.	+1 if both do not have another neighbor
		-1 if the prediction has another neighbor but the comparison not
		-1 if the comparison has another neighbor but the prediction not
		-2 if the prediction has 2 other neighbors but the comparison not
		-3 if the prediction has 3 other neighbors but the comparison not
		+1 if both have one more neighbor
		0 if the prediction has 1 and the comparison has 2 more neighbors
		-1 if the prediction has 1 and the comparison has 3 more neighbors

2nd side score	Compares the neighbors for each segment, very strict both atom and bond must match. Can only match ones.	+1 if matched

Table S3. Explanation of the scoring system. The paths/matrices are walked through and compared to each other. The main focus of this point system is the separation between good and bad matches for a given prediction. The last three scores are combined in the server output, the resulting ‘side chain score’ can be interpreted as a general evaluation of the given side chain.